

JPAB

CLIPPEDIMAGE= JP410063679A

PAT-NO: JP410063679A

DOCUMENT-IDENTIFIER: JP 10063679 A

TITLE: INFORMATION PRESENTATION DEVICE

PUBN-DATE: March 6, 1998

INVENTOR-INFORMATION:

NAME

FUJIKAWA, EIJI

KUBOSONO, RIYUUJI

YADA, KENICHI

ASSIGNEE-INFORMATION:

NAME

COUNTRY

NIPPON TELEGR & TELEPH CORP <NTT>

N/A

APPL-NO: JP08222222

APPL-DATE: August 23, 1996

INT-CL (IPC): G06F017/30

ABSTRACT:

PROBLEM TO BE SOLVED: To present the information of a field in which a user is interested usefully and completely as much as possible by grasping the range of user's interest in the case of presenting information.

SOLUTION: Weight (weight A) related to a specific keyword applied in order to converge information out of keywords included in a profile 5 and another weight (weight B) using relation among keywords in order to grasp user's interest in a wide range are calculated and the suitable score of information presented to a user is calculated by properly using the two kinds of weight A, B, so that the range of user's interest can be estimated and information can be presented.

COPYRIGHT: (C)1998, JPO

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-63679

(43) 公開日 平成10年(1998) 3月6日

(51) Int.Cl.⁶

G 0 6 F 17/30

識別記号

庁内整理番号

F I

G 0 6 F 15/403

技術表示箇所

3 5 0 C

3 4 0 A

審査請求 未請求 請求項の数 2 O L (全 9 頁)

(21) 出願番号

特願平8-222222

(22) 出願日

平成8年(1996) 8月23日

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 藤川 英士

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72) 発明者 窪園 竜二

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72) 発明者 矢田 健一

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

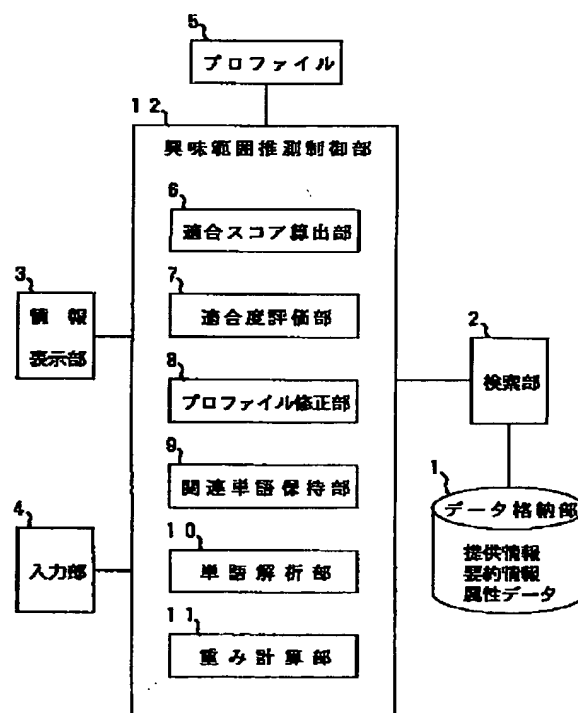
(74) 代理人 弁理士 小笠原 吉義 (外1名)

(54) 【発明の名称】 情報提供装置

(57) 【要約】

【課題】 情報を提供する場合に、利用者の興味の及ぶ範囲を捉えて、利用者の興味ある分野の情報をできるだけ無駄なく、漏れなく提供する。

【解決手段】 プロファイル5の中のキーワードに関して、情報を絞り込むために付与した特定キーワードに関する重み(重みA)の他に、利用者の興味を広い範囲で捉えるためにキーワード間の関連度を用いて別の重み(重みB)を算出し、2種類の重みを使い分けて利用者に対する提供情報の適合スコアを算出することにより、利用者の興味の及ぶ範囲を推測して情報を提供する。



【特許請求の範囲】

【請求項1】 利用者が検索キーを入力しなくても、データベースに追加された情報の中から利用者の興味や関心事に関する情報を自動的に検索して提供する情報提供装置において、利用者に提供する提供情報に付与されたキーワードとキーワードの重みを属性データとして保持する手段と、利用者が興味をもっている対象とその興味の度合いを表すものとして、キーワードとキーワードの第1の重みを保持する手段であるプロファイルと、指定したキーワードについて意味的な関連度を保持する手段と、前記プロファイル中のキーワードに対して単語が広い意味を持つ単語であるか狭い意味を持つ単語であるかを判断し、その判断結果から前記意味的な関連度を保持する手段で保持している関連度を参照して利用者の興味を広い範囲で捉えるための第2の重みを計算する手段と、前記提供情報の属性データから得られる重みと前記プロファイル中のキーワードに関する前記第1および第2の重みを用いて、提供情報が利用者の興味や関心事とどの程度合致しているかの指標である適合スコアを算出し、算出した適合スコアに基づいて情報を提供する手段と、利用者に対する提供情報の適合度を評価する手段と、適合度を参照して前記プロファイルに登録されているキーワードの追加・削除およびキーワードの重みを変化させる手段とを有し、前記第1の重みと前記第2の重みを使い分けることにより、利用者の興味及んでいる範囲を推測して情報を提供するようにしたことを特徴とする情報提供装置。

【請求項2】 利用者が検索キーを入力しなくても、データベースに追加された情報の中から利用者の興味や関心事に関する情報を自動的に検索して提供する情報提供装置において、利用者に提供する提供情報に付与されたキーワードとキーワードの重みを属性データとして保持する手段と、利用者が興味をもっている対象とその興味の度合いを表すものとして、キーワードとキーワードの第1の重みを保持する手段であるプロファイルと、利用者が獲得した情報に付与されたキーワードについて、2つのキーワードが同じ情報中に出現する割合である共起関係を保持する手段と、該キーワード共起関係を用いて、キーワード間の関連度を算出する手段と、キーワード間の関連度を保持する手段と、前記プロファイル中のキーワードに対して単語が広い意味を持つ単語であるか狭い意味を持つ単語であるかを判断し、その判断結果から前記キーワード間の関連度を保持する手段で保持している関連度を参照して利用者の興味を広い範囲で捉えるための第2の重みを計算する手段と、前記提供情報の属性データから得られる重みと前記プロファイル中のキーワードに関する前記第1および第2の重みを用いて、提供情報が利用者の興味や関心事とどの程度合致しているかの指標である適合スコアを算出し、算出した適合スコアに基づいて情報を提供する手段と、利用者に対する提

供情報の適合度を評価する手段と、適合度を参照して前記プロファイルに登録されているキーワードの追加・削除およびキーワードの重みを変化させる手段とを有し、前記第1の重みと前記第2の重みを使い分けることにより、利用者の興味及んでいる範囲を推測して情報を提供するようにしたことを特徴とする情報提供装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はデータベースに蓄積された情報の中から、利用者による検索キーの入力を必要とせず、利用者が興味を持っている情報を提供するための情報提供装置に関するものである。

【0002】

【従来の技術】従来の利用者が検索キーを入力しなくても、データベースに追加された情報の中から利用者の興味や関心事に関する情報を検索して提供する情報提供システムにおいては、以下のような方法がとられていた。

【0003】(1)プロファイルと提供情報の属性データに保持されたキーワードとキーワードの重みを用いて利用者に対する提供情報の適合スコアを算出し、適合スコアの高い情報を提供する。

【0004】この際に、提供情報の属性データ中のキーワードの重みを決定する方法として、個々の提供情報の特徴を捉えるために、特定キーワード(固有名詞などの意味の狭いキーワード)の重みを大きくし、また、プロファイル中のキーワードの重みを決定する方法として、利用者がこれまでに獲得した提供情報の属性データであるキーワードの重みをそのまま用い、利用者提供する情報を特定キーワードに関する狭い範囲に絞り込んで提供していた。

【0005】(2)一方、プロファイルに保持されている各キーワードに関して、シソーラス(単語の上位語・下位語を記した辞書)を用いて上位語を辿り、導き出された単語で検索することによって、利用者の興味を広く捉えて情報を提供しようとする手法もとられていた。

【0006】

【発明が解決しようとする課題】上記従来の情報提供装置においては、以下の問題があった。

(1)個々の提供情報の特徴を捉えるために特定キーワードの重みを大きくした場合、提供する情報を特定キーワードに関する狭い範囲に絞り込むため、利用者の興味に対する一部の情報しか提供できないという欠点があった。

【0007】例えば、分野 α に興味のある人のプロファイルに、分野 α に関連する固有名詞のキーワード β 、 γ 、 δ が登録されていたとする。従来の方法によれば β 、 γ 、 δ は重みが大きく、 β 、 γ 、 δ による検索では、分野 α に関する β 、 γ 、 δ に関連する情報しか提供できない。したがって、利用者が分野 α 全般に興味を持っている場合に、利用者の興味に対して漏れなく情報を

提供できないという欠点があった。

【0008】(2) プロファイル中のキーワードに関して、シソーラスを用いてその上位語を辿って検索することによって、利用者の興味を広く捉えようとする手法では、あらかじめ検索に用いるキーワードをシソーラスに登録しておく必要がある。

【0009】また、シソーラス自体が単に単語の上位語・下位語を導き出すためのものであるため、プロファイル中のキーワードからどの程度のシソーラス上の上位語を辿っていけばよいか分からず、上位語を辿りすぎると提供情報の範囲が広くなり過ぎて、無駄な情報を提供してしまうという欠点があった。

【0010】本発明の目的は、プロファイル中のキーワードに関して、これまでの特定キーワードに関する狭い範囲に絞って情報を提供するために付与された重み(重みA)の他に、利用者の興味をより広く捉えて情報を提供するために、キーワードの共起関係等から求まるキーワード間の関連度を用いて別の重み(重みB)を算出し、2種類の重みを使い分けて利用者に対する提供情報の適合スコアを算出し、利用者の興味及び範囲を推測して、利用者の興味ある分野の情報をできるだけ無駄なく、漏れなく提供することにある。

【0011】上記の例でいえば、利用者の興味が分野 α に関するキーワード β , γ , δ にしか向けられていないのか、分野 α 全般に向けられているのかを把握し、もし利用者の興味が分野 α 全般に向けられているとすれば、分野 α に関する情報の中でキーワード β , γ , δ に関する情報以外の情報を提供し、利用者の興味がキーワード β , γ , δ にしか向けられていなければ、分野 α に関する情報の中からキーワード β , γ , δ に関する情報のみを提供する。

【0012】

【課題を解決するための手段】上記課題を解決するため、本発明は以下のように構成される。

(1) 提供情報の属性データを蓄積する属性データ蓄積部と、利用者が興味をもっている対象とその興味の度合いを表すものとして、キーワードとキーワードの重みを保持する手段であるプロファイルと、該プロファイルと前記属性データを用いて、提供情報が利用者の興味や関心事とどの程度合致しているかの指標である適合スコアを算出する適合スコア算出部と、該適合スコア算出部の結果に基づいて出力部に提供情報を出力したときに、利用者の提供情報に対する適合度を評価する適合度評価部と、適合度を参照して、プロファイルに保持されているキーワードの追加・削除およびキーワードの重みを変化させるプロファイル修正部とから構成される従来の情報提供装置に加え、プロファイルに登録されたキーワードについて関連したキーワードとその関連度を保持する関連単語保持部と、指定した単語(キーワード)について意味的に広い単語であるか、意味的に狭い単語であるか

を判断する単語解析部と、利用者の興味の対象を広く捉えて情報を提供するためにプロファイル中のキーワードの重みBを計算する重み計算部と、重み計算部によって計算した重みBを用いて情報を提供したときに、利用者の興味及び範囲を推測して情報を提供する制御を行う興味範囲推測制御部とを備える。

(2) または、上記の関連単語保持部の代わりに、利用者が獲得した情報に付与されたキーワードについて、2つのキーワードが同じ情報中に出現する割合である共起関係を保持する共起関係保持部と、該キーワード共起関係を用いて、キーワード間の関連度を算出する関連度算出部と、キーワード間の関連度を保持する関連単語保持部とを備える。

【0013】本発明の作用は、以下のとおりである。上記(1)の発明では、検索を行う際、普通名詞のような広い意味を持つ単語を単独に用いても特定の分野の情報を得ることは難しいが、広い意味を持つ単語を複数組み合わせ合わせて検索を行うことによって、特定の分野の情報を得やすいという性質を用いることにより、利用者が興味を持っている具体的な対象(固有名詞のような意味の狭いキーワード)に関連があり、かつ、利用者が興味を持っている対象を広い範囲で捉えて情報を提供できる。

【0014】複数の意味の広い単語を用いて検索して得られた情報を利用者に提供し、利用者の提供情報に対する取り扱い方などから利用者に対する情報の適合度を評価する。ここで、適合度が低く評価された場合には、利用者の興味は特定の狭い意味のキーワードに関する情報に限られたものと判断できる。一方、適合度が高く評価された場合には、利用者の興味はその広い意味を持つ単語のキーワードを含んだ分野全体に及んでいる可能性がある」と判断できる。

【0015】上記(2)の発明では、利用者の興味を広く捉えるために、利用者の情報獲得の履歴からプロファイル中に登録されたキーワードの関連単語とそれらの関連度を蓄積していくことにより、シソーラス等の類義語辞書や関連語辞書を使用することなく、前述した(1)の発明でのキーワード間の関連度を保持することを実現することができる。

【0016】従って、本発明を用いることにより、最終的に利用者の興味及び範囲を捉えて、利用者が興味を持っている対象に関する情報をできるだけ無駄なく、漏れなく提供することができる。

【0017】

【発明の実施の形態】

(第1の実施の形態) 図1は本発明を実施する装置構成の一形態を示すブロック図である。

【0018】本発明を実施するための装置は、例えば図1に示すように、テキスト、音声、画像などから構成される提供情報、その要約情報、およびその属性データを格納するデータ格納部1と、蓄積された情報に対して検

5

索を行う検索部2と、提供情報を出力する情報表示部3と、利用者の装置に対する操作を受け付ける入力部4と、利用者が過去に獲得した提供情報の属性データを参照してキーワードとその重みを保持するプロファイル5と、このプロファイル5とデータ格納部1に格納された情報の属性データから適合スコアを算出する適合スコア算出部6と、利用者の提供情報に対する取り扱い方や利用者の提供情報に対する評価などから利用者の情報に対する適合度を評価する適合度評価部7と、適合度評価部7で評価した適合度を参照して、プロファイル5に保持されているキーワードとキーワードの重みを修正するプロファイル修正部8から構成される従来の情報提供装置に加えて、関連単語保持部9、単語解析部10、重み計算部11等を持つ。

【0019】関連単語保持部9は、あらかじめ登録された単語間の関連度を保持しているものである。単語解析部10は、プロファイル5中のキーワードが広い意味を持つ単語か、狭い意味を持つ単語かを判別するために、プロファイル5中のキーワードが例えば普通名詞か固有

名詞かを判別する。

【0020】重み計算部11は、利用者の興味の対象を広く捉えて情報を提供するために、単語解析部10の判別結果から関連単語保持部9を参照してプロファイル5の重みBを計算する。

【0021】興味範囲推測制御部12は、重み計算部11によって計算した重みBを用いて利用者の興味が及んでいる範囲を推測した情報提供の制御を行う。図2は、データ格納部1での提供情報とその要約情報とその属性データの格納例を示す。データ格納部1には、提供情報の本体に加えて、その情報の要約と属性データとが、図2に示すように格納される。属性データは、キーワードとその重みである。

【0022】図3は、プロファイル5でのキーワードとその重みの格納例を示す。プロファイル5には、各キーワードに対して、第1の重みAと第2の重みBが、例えば図3に示すように格納される。なお、重みBは、情報提供時に重みAから動的に算出されるので、必ずしもあらかじめプロファイル5中に格納されていなくてもよい。

【0023】図4は、意味的な関連度の説明図である。関連単語保持部9が保持する関連単語の意味的な関連度は、あるキーワードと他のキーワードとがどれだけ意味的に近いかを数値で表したもので、例えばn個のキーワード $\{k_1, k_2, \dots, k_n\}$ について、キーワード k_i とキーワード k_j との関連度を w_{ij} とおくと、図4に示すような対称行列として表現される。ただし、 $w_{ij} = w_{ji}$ である。

【0024】図5は本装置の処理の基本フローである。図5において、まず処理P1では、興味範囲推測制御部12によって提供する情報の総数を決定する。提供する

6

情報の総数は、利用者に総数を入力させる場合や、あらかじめ固定の数とする場合などがある。

【0025】処理P2では、興味範囲推測制御部12によって重みAで情報を提供する割合と重みBで提供する割合を決定する。重みA/重みBの設定方法については後述する。

【0026】重みAを用いた情報の提供では、処理P3により、適合スコア算出部6によって、プロファイル5に保持されたキーワードと属性データのキーワードとを照合し、合致するキーワードのそれぞれの重み（プロファイル5のキーワードの重みは重みAを用いる）を掛け合わせて、それらの和をとることにより、適合スコアを算出する。

【0027】また、重みBを用いた情報の提供では、処理P4によって、重み計算部11により重みAから重みBを計算する。この計算は、例えば次のように行う。プロファイル5中のキーワードについて、m個の固有名詞のキーワード(P_1, P_2, \dots, P_m)と、n個の普通名詞のキーワード(Q_1, Q_2, \dots, Q_n)からなる場合、固有名詞の重みAを(p_1, p_2, \dots, p_m)とし、普通名詞の重みAを(q_1, q_2, \dots, q_n)とする。

【0028】固有名詞 P_i と普通名詞 Q_j の関連度を R_{ij} とし、固有名詞 P_i および普通名詞 Q_j の重みBをそれぞれ x_i, y_j とすると、 x_i, y_j は次式の計算で求められる。

【0029】

$$x_i = 0 \quad (i = 1, 2, \dots, m)$$

$$y_j = q_j \sum R_{kj} p_k \quad (j = 1, 2, \dots, n)$$

〔ただし、 Σ は $k = 1$ から m までの総和を表す〕

次に、処理P5により、適合スコア算出部6によって、プロファイル5に保持されたキーワードと属性データのキーワードとを照合し、合致するキーワードのそれぞれの重み（プロファイル5のキーワードの重みは重みBを用いる）を掛け合わせて、それらの和をとることにより、適合スコアを算出する。

【0030】処理P6では、処理P3および処理P5で算出された適合スコアに基づいて、要約情報を情報表示部3に提示する。次に処理P7では、利用者に提供情報の評価を直接的に求める。あるいは図6に示すように、獲得した情報の内容や、情報の閲覧時間などを定量的に測り、適合度の評価を行う。すなわち、図6に示すように、利用者が要約情報だけでなく提供情報の本体を要求して獲得した場合には、利用者の適合度が大きいと評価する。また、情報の閲覧時間が長いほど、利用者の適合度が大きいと評価する。情報の閲覧時間は、例えば提供情報を送信してから、次に利用者から何らかの要求があるまでの時間によって計測することができる。

【0031】処理P8では、適合度評価部7によって評価された結果をもとに、図7に示す処理により、プロフ

10

20

30

40

50

ファイル5の修正を行う。処理P9では、情報提供を続けるかどうかを判断し、情報提供を続ける場合には、処理P1へ戻って同様に処理を繰り返す。情報提供を続けない場合には、処理を終了する。

【0032】図7は、図5に示す処理P8におけるプロファイル5の修正フローを示す。重みBは、重みAに基づいて毎回計算されるので必ずしも保存する必要はなく、プロファイル5中の重みの修正は、重みAに対してだけ行われる。ただし、重みBによって適合スコアを計算して提供した情報については、適合スコアを計算する際に「普通名詞」しか考慮していないので、プロファイル中の該当キーワードのうち、普通名詞のキーワードの重みAのみを変化させる。

【0033】図7の処理P21では、利用者に対する適合度が大きいかどうかを判別する。利用者に対する適合度が大きい場合、処理P22へ進み、小さい場合、処理P27へ進む。

【0034】処理P22では、データ格納部1から得た属性データのキーワードで、プロファイル5を検索する。次の処理P23による判定により、プロファイル5中に該当するキーワードが存在する場合、処理P24によって、プロファイル5中の該当キーワードの重みを上げる。該当するキーワードが存在しない場合、処理P25によって、プロファイル5中に該当キーワードとその重みを追加する。

【0035】次の処理P26では、評価対象の情報が他にあるかどうかを判定し、評価対象の情報がもうない場合には処理を終了する。評価対象の情報がある場合には処理P22以下の処理を同様に繰り返す。

【0036】一方、利用者に対する適合度が小さい場合、処理P27により、属性データのキーワードでプロファイル5を検索する。そして、処理P28により、プロファイル5中の該当キーワードの重みを下げる。次に処理P29の判定により、評価対象の情報がもうない場合には処理を終了し、ある場合には処理P27へ戻って同様に処理を繰り返す。

【0037】図5の処理P2等に関連する重みA/重みBの設定方法の例を、①提供する情報の総数を利用者が決定する場合と、②全ての情報を適合スコアの高い順にリストとして提供する場合について説明する。

【0038】① 提供する情報の総数を利用者が決定する場合

この場合には、まず、重みAで提供する情報と重みBで提供する情報の割合を保存しておく。そして、利用者へ情報を提供する際に、重みAで適合スコアを計算して提供する情報と、重みBで適合スコアを計算して提供する情報とを、保存している割合に従って混在した形で提供し、利用者の適合度から割合を修正する。

【0039】例えば、提供する情報の総数が10個であり、重みAと重みBでの提供情報の割合が7:3であっ

たとする。重みAで適合スコアを計算し、適合スコアが大きい順に7つ(7割)の情報を提供し、また重みBで適合スコアを計算し、適合スコアが大きい順に3つ(3割)の情報を提供する。

【0040】適合度を計算し、次のように計算した割合を次回に提供する際に用いる。例えば、次回に提供する情報の比率は、次式によって決定する。

重みAによる情報提供:重みBによる情報提供=(重みAによって提供した情報の適合スコアの和)/(重みAの比率):(重みBによって提供した情報の適合スコアの和)/(重みBの比率)

これらの計算は、興味範囲推測制御部12で行う。

【0041】② 全ての情報を適合スコアの高い順にリストとして提供する場合

この場合には、重みAで提供する情報と重みBで提供する情報の割合は決定しない。重みAで適合スコアを計算し、適合スコアが大きい順に要約情報をリストにして提示する部分と、重みBで適合スコアを計算し、適合スコアが大きい順に要約情報をリストにして提示する部分とを、同一画面上に配置し提供する。

【0042】以上のように、興味範囲推測制御部12は、重みBを用いて情報を提供し、利用者の適合度を評価した際に利用者の興味が分野全体に及ぶのか、特定キーワードのみに及ぶのかを推測して情報を提供する制御を行う。

【0043】本実施の形態のように、固有名詞に大きい重みを付与することにより狭い範囲に絞って情報を提供する方法と、プロファイル5中の固有名詞に対して関連度の高い普通名詞群に大きい重みを付与することにより利用者の興味の対象をより広く捉えて情報提供する方法とを組み合わせる情報提供を行い、そのときの利用者の適合度を評価することによってフィードバックを行い、利用者の興味の及ぶ範囲を特定することにより、利用者が興味を持っている対象に関する情報をできるだけ無駄なく、漏れなく提供することが可能になる。

【0044】〔第2の実施の形態〕次に、本発明の第2の実施の形態について説明する。図8は、データ格納部1に提供情報そのものを格納する代わりに、提供情報の所在を表すアドレス情報を格納する場合の第2の実施の形態の装置構成を示すブロック図である。

【0045】データ格納部1には、アドレス情報、要約情報、属性データを格納する。アドレス情報としては、例えばインターネットにおけるURL等がある。この実施の形態の場合、データ格納部1の中の属性データによって適合スコアを算出し、情報を提供するときは、データ格納部1の中のアドレス情報により通信部14を介して情報を取得する。

【0046】他は、前述した第1の実施の形態と同様である。

〔第3の実施の形態〕図9は、提供情報としてキーワー

ドとその重みの付与されていないテキスト情報を外部から格納する場合の第3の実施の形態の装置構成を示すブロック図である。

【0047】この実施の形態では、テキスト情報は、情報投入部15からデータ格納部1に投入される。情報解析部16によってキーワードが抽出され、キーワードの重みを付与されて、それらはデータ格納部1に格納される。

【0048】共起関係保持部17では、利用者が獲得した情報に付与されたキーワードの共起関係を保持する。共起関係は、複数の情報を扱う場合において2つのキーワードが同じ情報の中に現れる割合のことであり、ここでは、利用者が獲得した情報について2つのキーワードが共通に現れる割合である。

【0049】関連度算出部18では、利用者が獲得した情報におけるキーワードの共起関係を用いて、例えば下記の文献[1]のpp.63-66に記載された手法において、文献を提供情報に、類似度を関連度に置き換えることによってキーワード間の関連度を算出し、関連単語保持部9に保持する。

【0050】すなわち、プロフィール5中のキーワードの集合を、 $K = \{k_1, k_2, \dots, k_N\}$ とし、これらのうちいずれかが現れる利用者が獲得した情報の集合を、 $D = \{d_1, d_2, \dots, d_N\}$ とする。そうするとキーワード $k_u, k_v \in K$ の情報集合Dでの出現状況は、利用者が獲得した情報のベクトルを用いて、次のように表される。

【0051】 $k_u : v_{ku} = (\delta_1^u, \delta_2^u, \dots, \delta_i^u, \dots, \delta_N^u)$

$k_v : v_{kv} = (\delta_1^v, \delta_2^v, \dots, \delta_i^v, \dots, \delta_N^v)$

ただし、 $\delta_i^u (\delta_i^v)$ 、 $1 \leq i \leq N$ は、キーワード $k_u (k_v)$ が情報 d_i に含まれるときに1、それ以外の際に0となる。これから、キーワード間の意味的な関連度は、情報集合D中でそれらが共通して生じる利用者が獲得した情報の数に比例した値を与える関連測度 $s : K \times K \rightarrow [0, 1]$ を定めることで計算できる。この関連測度sの定式化は、例えば計算方法が簡単で応用範囲の広い集合論的測度やコサイン関数などのよく知られた手法を用いればよい。

【0052】以下、前述した第1の実施の形態と同様である。

参考文献[1]「情報検索」pp.63-66、伊藤哲郎、昭晃堂(1986)。

【0053】

【発明の効果】以上説明したように、本発明によれば、利用者が検索キーを入力することなしに、データベースに蓄積された情報の中から、利用者の興味の対象に関する情報を従来の装置に比べて無駄なく、漏れなく提供することができるようになる。

【図面の簡単な説明】

【図1】第1の実施の形態における装置構成を示すブロック図である。

【図2】第1および第3の実施の形態におけるデータ格納部での提供情報とその要約情報とその属性データの格納例を示す図である。

【図3】第1、第2および第3の実施の形態におけるプロフィールでのキーワードとその重みの格納例を示す図である。

【図4】意味的な関連度の説明図である。

【図5】第1、第2および第3の実施の形態における処理の基本フローである。

【図6】第1、第2および第3の実施の形態における適合スコアの算出の例を示す図である。

【図7】第1、第2および第3の実施の形態におけるプロフィールの修正のフローチャートである。

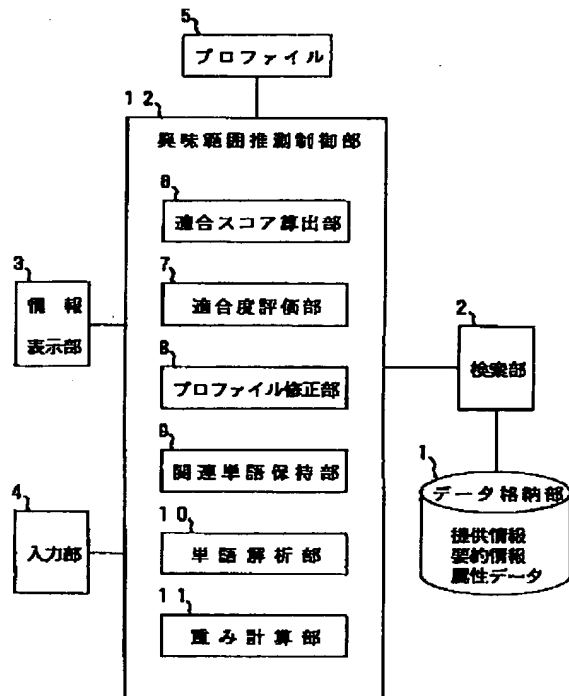
【図8】第2の実施の形態における装置構成を示すブロック図である。

【図9】第3の実施の形態における装置構成を示すブロック図である。

【符号の説明】

- 1 データ格納部
- 2 検索部
- 3 情報表示部
- 4 入力部
- 5 プロフィール
- 6 適合スコア算出部
- 7 適合度評価部
- 8 プロフィール修正部
- 9 関連単語保持部
- 10 単語解析部
- 11 重み計算部
- 12 興味範囲推測制御部
- 14 通信部
- 15 情報投入部
- 16 情報解析部
- 17 共起関係保持部
- 18 関連度算出部

【図1】



【図3】

キーワード	重みA	重みB
△△電信電話株式会社	0.7482	0
マルチメディア	0.4517	1.145
通信	0.1475	0.254
電話	0.1245	0.584
FAX	0.1154	0.751
⋮	⋮	⋮

【図2】

1. 情報の本体:

情報の要約:

情報の属性データ:

キーワード	重み
△△電信電話株式会社	1.2543
マルチメディア	0.8425
ネットワーク	0.7848
コンピュータ	0.5245
インターネット	0.5154
⋮	⋮

【図4】

意味的な関連度

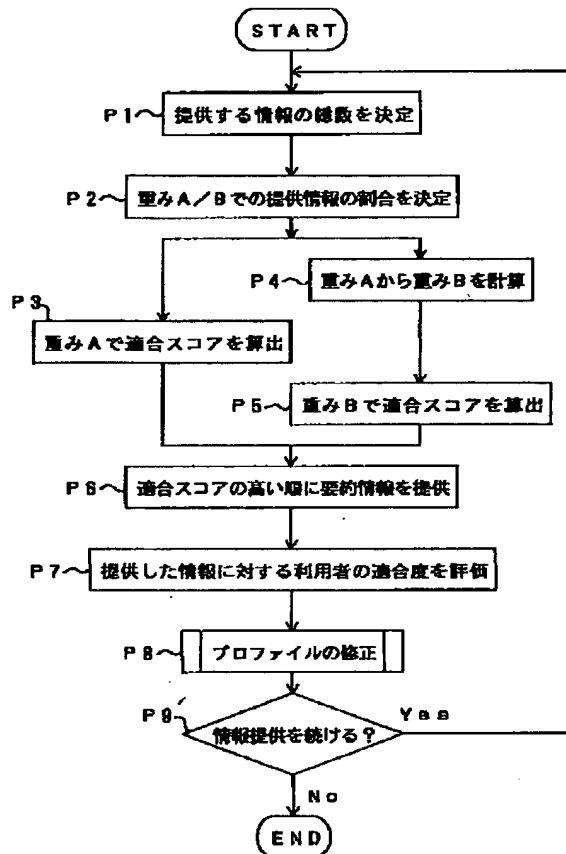
	k_1	k_2	k_3	...	k_n
k_1		w_{12}	w_{13}	...	w_{1n}
k_2	w_{21}		w_{23}	...	w_{2n}
k_3	w_{31}	w_{32}		...	w_{3n}
⋮	⋮	⋮	⋮	⋮	⋮
k_n	w_{n1}	w_{n2}	w_{n3}	...	

 k_i : キーワード w_{ij} : 関連度

【図6】

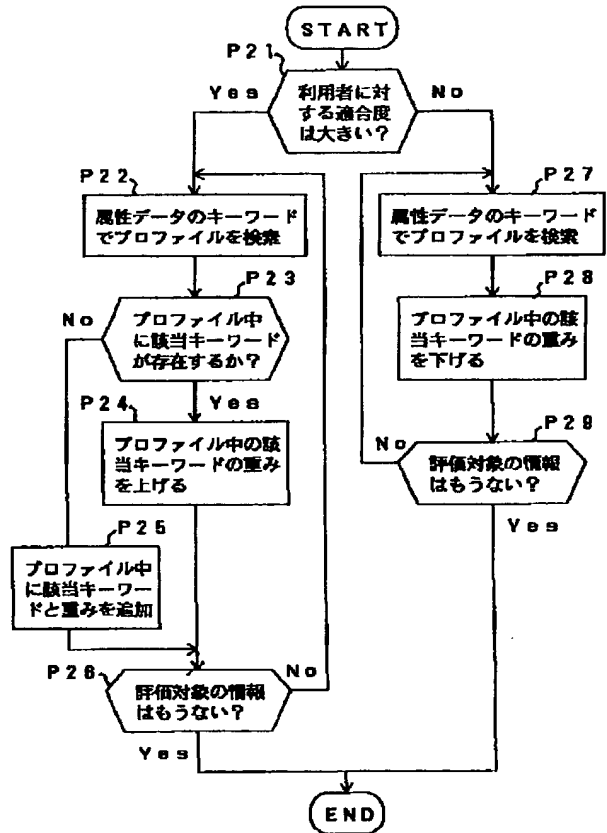
獲得した 情報の内容	要約情報 → 提供情報の本体 選択
情報の閲覧時間	短い → 長い
利用者の適合度	小さい → 大きい

【図5】

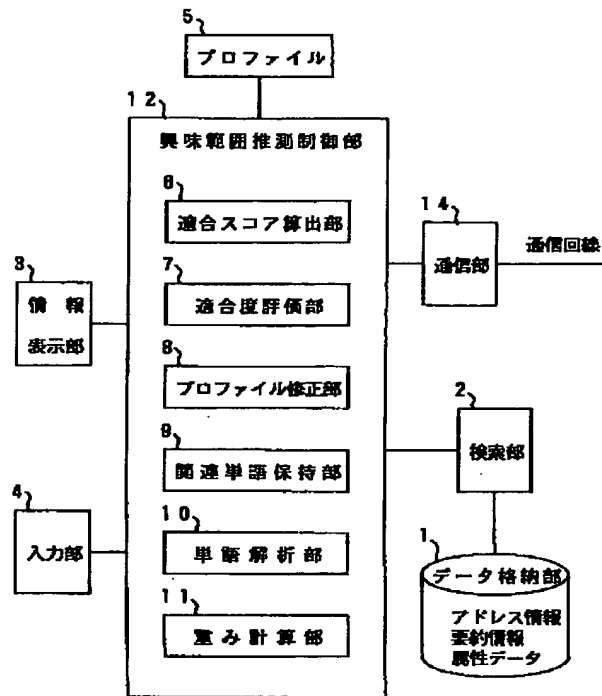


【図7】

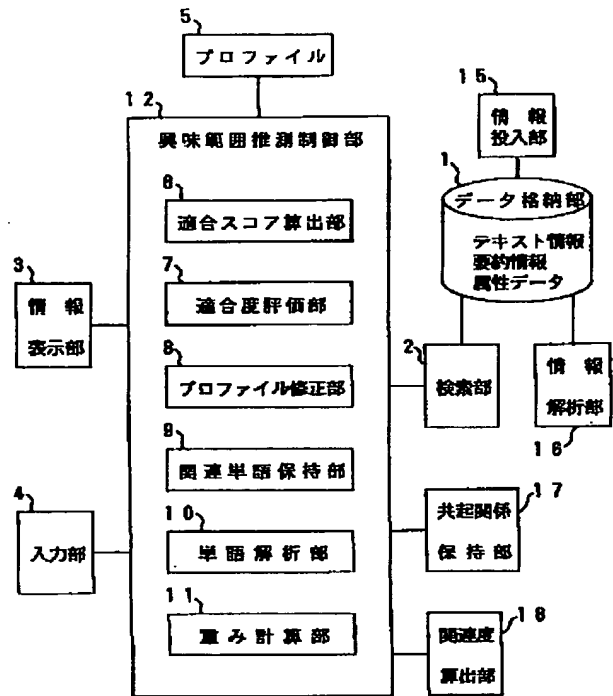
プロフィール修正フロー



【図8】



【図9】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.